(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 828 004 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
11.03.1998 Bulletin 1998/11

(51) Int Cl.⁶: C12N 15/61, C12N 15/81, C12N 9/90, C12N 1/19

(21) Application number: 97306871.1

(22) Date of filing: 04.09.1997

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: 04.09.1996 JP 234287/96

(71) Applicant: SUNTORY LIMITED
Kita-ku, Osaka-shi, Osaka (JP)

(72) Inventors:
- Sakai, Yasuyoshi
  Otsu-shi, Shiga (JP)
- Kato, Nobuo
  Kameoka-shi, Kyoto (JP)
- Shibano, Yuji
  Toyonaka-shi, Osaka (JP)

(74) Representative: Stoner, Gerard Patrick et al
MEWBURN ELLIS
York House
23 Kingsway
London WC2B 6HP (GB)

Remarks:
The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) Protein disulfide isomerase gene derived from strain of methylotrophic yeast

(57) A protein derived from a strain of methylotrophic yeast which has a protein disulfide isomerase activity having the amino acid sequence as set forth in SEQ ID No. 1, or protein in which said amino acid sequence has been modified by deletion or addition of one or a few amino acids, or substitution with other amino acid(s) and which has a protein disulfide isomerase activity, and a process for production thereof.

EP 0 828 004 A2

Printed by Jouve, 75001 PARIS (FR)

Description

The present invention relates to protein disulfide isomerase, an enzyme which promotes formation of protein conformation by catalyzing formation of disulfide bonds in a protein, and to a gene thereof. The present invention relates, among the protein disulfide isomerases, to protein disulfide isomerase derived from a strain of methylotrophic yeast, a microorganism suitable for industrial production of valuable proteins due to its high efficiency of expression of heterologous genes and secretion of the expression products, and to a gene thereof.

Protein disulfide isomerase (PDI) is a major protein present in the lumen of the endoplasmic reticulum (referred to hereinafter as ER) and it was first discovered as having an activity which effects oxidative refolding of a reduced RNase (Goldberger, R.F. et al. (1963) J. Biol. Chem. 238: 628-635). PDI is believed to be an enzyme which catalyzes formation of stable conformation by recombining disulfide bonds of secretory proteins.

It has been pointed out that, in the case of heterologous proteins, in particular secretory proteins, which often have disulfide bonds, recombination of disulfide bonds by PDI as well as protein folding by peptidyl-prolyl-cis-trans isomerase (PPI) represent the rate-limiting step in the secretory process of proteins (Gething, M.J. and Sambrook, J. (1992) Nature 355: 33-45). It has also been demonstrated that PDI promotes folding of proteins consisting of a single domain such as RNase in vitro as well (Jaenicke, R. (1993) Curr. Opin. Struct. Biol. 3: 104-102).

On the other hand, because strain of methylotrophic yeasts grow using methanol as the sole carbon source and they have high yields of cells, they have been used for the production of materials for use in the synthetic chemical industry including, for example, aldehydes such as formaldehyde, epoxides, methylethylketone, and formic acid. Research has been conducted on the possible utilization of the cells per se as a protein source, and the utilization for production of cell components such as amino acids, vitamins, and the like, and some have been put into practical use. In recent years, furthermore, an expression system of heterologous genes using strain of methylotrophic yeasts as the host has been developed and it has been shown that said system has a higher productivity than Saccharomyces yeasts (Japanese Unexamined Patent Publication (Kokai) No. 5-344895).

Its productivity is high especially for secretory proteins. For example, the productivity of glucoamylase derived from filamentous fungi of the genus Rhizopus was 3.4 g/l, which is about 10 times higher than the productivity by Saccharomyces yeasts (Sakai, Y., et al. (1998) Biochim. Biophys. Acta 1308: 81-87). As the strain of methylotrophic yeast, there are known Candida boidinii, Pichia pastoris, Hansenula polymorpha, and the like.

When heterologous proteins are produced by secretory production using recombinant DNA technology, the efficiency of the secretion is thought to be increased by enhancing the speed of folding proteins. Based on such an idea, an example has been disclosed in which the amount secreted of human albumin was increased by about 60% on the average by coexpressing a human PDI gene with the desired gene in a Saccharomyces yeast (Japanese Unexamined Patent Publication (Kokai) No. 5-36771).

Formation or exchange of disulfide bonds which are necessary for appropriate folding of proteins requires environments suitable therefor. For that purpose, eukaryotic cells have intracellular compartments such as the ER or the Golgi apparatus, etc. While passing through the compartments, secretory proteins are subjected to suitable folding or addition of sugar chains and then are secreted out of the cell by means of exocytosis. Many of the secretory proteins of eukaryotic origin have intramolecular disulfide bonds, and formation and exchange of these disulfide bonds taking place in the ER are essential for formation of protein conformation and its secretion.

Accordingly, the PDI which catalyzes reactions for formation and/or exchange of disulfide bonds must be localized or stay in the ER. For this purpose the PDI has a unique amino acid sequence called an ER retention signal sequence at the C-terminal. As ER retention signal sequences there are known Lys-Asp-Glu-Leu (SEQ ID No. 2) for animals and His-Asp-Glu-Leu (SEQ ID No. 3) for Saccharomyces yeasts. When the human PDI gene as described above was expressed in a Saccharomyces yeast, the ER retention signal sequence of the human PDI did not fully function, which was possibly due to inadequate localization of the PDI in the ER. Thus, it is believed that even the highly expressed PDI gene did not cause enhancement in the PDI activity commensurate with the expression in the ER, and accordingly the increment of the amount secreted of the coexpressed secretory protein remained at a value of 60%.

In order for the PDI expressed in a strain of methylotrophic yeast to fully perform its functions, it is preferred to use the PDI derived from a strain of methylotrophic yeast. The reason why the strain of methylotrophic yeast has a high ability of secreting protein as described above is that recombination of disulfide bonds by the PDI which is the rate-limiting step of the protein secretion process takes place efficiently and that the PDI derived from the strain of methylotrophic yeast has a higher specific activity than the PDI derived from other sources or has a higher activity in the ER. However, the PDI of the strain of methylotrophic yeast or the gene thereof was unknown. Accordingly, no studies had been carried out on enhancement of productivity in the expression system of the strain of methylotrophic yeast by using the above PDI or the gene thereof.

The inventors have carried out intensive studies to clone the PDI gene carried by strain of methylotrophic yeast, to elucidate the nucleotide sequence thereof, and to reveal the characteristics of the PDI of strain of methylotrophic yeast. Thus, it is the object of the present invention to provide the PDI gene derived from strain of methylotrophic yeast

in order to effect secretory production of heterologous genes by strain of methylotrophic yeast in a more efficient manner.

In order to attain the above-mentioned objective, the inventors have obtained a DNA fragment amplified by the PCR using as a primer an oligonucleotide synthesized based on an amino acid sequence of the conserved region present in the active site of the PDI. By means of the colony hybridization method using this amplified DNA fragment as a probe, the inventors have cloned the PDI gene of the strain of methylotrophic yeast Candida boidinii, and demonstrated the nucleotide sequence of said gene and the amino acid sequence of said PDI. Furthermore, by coexpressing the peroxidase gene derived from a filamentous fungus in the strain of methylotrophic yeast transformed with said PDI gene, the inventors have successfully increased by about 10 times the amount secreted of said peroxidase and have accomplished the present invention.

Thus, the present invention provides a protein derived from a strain of methylotrophic yeast which has a protein disulfide isomerase activity having the amino acid sequence as set forth in SEQ ID No. 1, or protein in which said amino acid sequence has been modified by deletion or addition of one or a few amino acids, or substitution with other amino acid(s) and which has a protein disulfide isomerase activity. The present invention also provides a gene encoding the PDI, a vector comprising said gene, and a host transformed with said vector, as well as a process for secreting in large amounts the desired protein by coexpressing the gene for said desired protein in said transformed yeast host.

BRIEF EXPLANATION OF THE DRAWINGS

Fig. 1 is a drawing that shows a restriction enzyme map of the 6.2 kb DNA fragment containing the PDI1 gene of Candida boidinii, the region for which the nucleotide sequence was determined, and the position and direction of the PDI1 gene thereof.

Fig. 2 is a drawing that shows the result of Southern hybridization demonstrating the presence of the PDI gene in C. boidinii.

Fig. 3 (a) is a drawing that shows the construction of the expression vector pNPO3 of the ARP gene and (b) is a drawing that shows the construction of the expression vector pNRPD of the PDI1 gene.

Fig. 4 is a drawing that shows a procedure of construction of the expression vector pNRPD for the PDI1 gene.

Fig. 5 is a schematic diagram showing the state in which the ARP gene has been integrated into the genomic DNA of the BPO17 strain, and a drawing that shows the result of Southern hybridization confirming it.

Fig. 6 is a schematic diagram showing the state in which the PDI1 gene has been integrated into the genomic DNA of the BPP1 strain, and a drawing that shows the result of Southern hybridization confirming it.

Fig. 7 is a drawing that shows the result of Northern analysis which analyzed the amount of the expressed PDI1 gene.

Fig. 8 is a drawing that shows the ARP activity in the culture liquid of each of the cultured BPO17 strain, the BPP1 strain, and the BUL strain.

DETAILED DESCRIPTION

The present invention is now explained in detail below.

First, the sequence, Cys-Gly-His-Cys, which is conserved in PDI's from a variety of sources as the active center of the exchange reaction of disulfide bonds was found at two sites in the amino acid sequence of the PDI derived from Saccharomyces cerevisiae. Based on the amino acid sequence of PDI of S. cerevisiae comprising said sequence, various primers for the PCR were designed with reference to the frequency of use of codons from the strain of methylotrophic yeast. Using these primers PCR reactions were carried out using the genomic DNA of the strain of methylotrophic yeast as a template, and the amino acid sequence deduced from the nucleotide sequence of the PCR reaction product thus obtained was confirmed to be analogous to the amino acid sequence of PDI of S. cerevisiae.

The genomic DNA of a strain of methylotrophic yeast is completely digested with various restriction enzymes and is fractionated on agarose gel electrophoresis. Using the above-mentioned PCR products as a probe, Southern hybridization is carried out to find a restriction enzyme which gives the smallest DNA fragment containing the entire region of the PDI gene. Using the genomic DNA of the strain of methylotrophic yeast which has been completely digested with the restriction enzyme, a genomic library is created, which is then subjected to colony hybridization using the above-mentioned PCR product as a probe to select clones having the PDI gene.

Plasmid is extracted from the selected clone, and is subjected to Southern hybridization to confirm that the plasmid contains the sequence of the above-mentioned PCR product. Furthermore, a restriction map of the inserted fragments of this plasmid is created, based on which subcloning is conducted to obtain the smallest DNA fragment containing the PDI gene. The nucleotide sequence of the DNA fragment obtained is determined and the amino acid sequence of the PDI derived from the strain of methylotrophic yeast is analyzed.

The PDI gene thus obtained derived from the strain of methylotrophic yeast can be highly expressed in the strain of methylotrophic yeast to prepare the PDI. As an expression vector for the PDI gene known vectors may be used.

and as an expression vector for the strain of methylotrophic yeast Candida boidinii, pNOTe1 or pTRex as described in Japanese Unexamined Patent Publication (Kokai) No. 5-344895 may be used. As the method for transforming the strain of methylotrophic yeast and the method for obtaining a transformant in which a foreign gene has been integrated into the chromosomal DNA thereof, a known method (Sakai, Y. et al. (1991), J. Bacteriol. 173: 7458-7463) can be used. Furthermore, the amount secreted of the desired secretory protein can be enhanced by coexpressing the PDI gene derived from the strain of methylotrophic yeast with the gene of the desired secretory protein in the strain of methylotrophic yeast.

Although the PDI derived from the strain of methylotrophic yeast had the ER retention signal Arg-Asp-Glu-Leu (SEQ ID No. 9) which is different from His-Asp-Glu-Leu (SEQ ID No. 3) derived from a Saccharomyces yeast, there is no doubt that the cells of C. boidinii recognize the former sequence which is of its own and the PDI is retained by the ER to fully perform its function. As expression vectors employed for expression of the PDI, those in which auxotrophic markers such as the above-mentioned pNOTe1 and pTRex have been replaced with the genes different from the ones used for the expression vector of the desired protein may be used. Furthermore, by imparting to the host strain of methylotrophic yeast auxotrophy corresponding to the two markers of the expression vector, transformation is possible by the method as described above. As the method for imparting auxotrophy to the strain of methylotrophic yeast, a known method (Sakai, Y. et al. (1991), J. Bacteriol. 173: 7458-7463) can be used.

EXAMPLES

The invention will be understood more readily with reference to the following examples; however these examples are intended to illustrate the invention and are not to be construed to limit the scope of the invention.

Example 1.

From Candida boidinii strain S2 (Tani, Y., et al. (1985) Agric. Biol. Chem. 49: 2699-2706), the PDI gene was obtained and the nucleotide sequence thereof was determined. Incidentally, said strain has been designated Candida boidinii SAM1958 and deposited as an international deposition under the Budapest Treaty on February 25, 1992, with National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, MITI, 1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki, 305, Japan, with the accession number FERM BP-3766.

(1) Amplification by PCR

Two sequences were noted as the amino acid sequences in the PDI of S. cerevisiae relating to Cys-Gly-His-Cys (SEQ ID No. 4), a sequence which is conserved in the PDI of various origins as the active center of the disulfide bond exchange reaction:

Pro-Trp-Cys-Gly-His-Cys-Lys (SEQ ID No. 5)

(amino acids No. 89 through No. 65 in the amino acid sequence of the PDI of Saccharomyces yeast),

Tyr-Ala-Pro-Trp-Cys-Gly-His (SEQ ID No. 6)

(amino acids No. 402 through No. 408 in the amino acid sequence of the PDI of Saccharomyces yeast).

Referring to the frequency of use of C. boidinii codons, oligonucleotide having the following nucleotide sequence corresponding to the amino acid sequence was synthesized.

That is, as the sense primer,

5'-CCGGAATTC CCT(A) TGG TGT(C) CGT(A) CAT(C) TGT(C) AA-3'
(SEQ ID No. 7),

and as the antisense primer,

5'-CGCGGATCC TG A(T)CC A(G)CA CCA A(T)GG A(G/T)GC A(G)T-3' (SEQ ID No. 8)

were synthesized. These oligonucleotides have on their 5'-end the sequence which recognizes EcoRI and BamHI, respectively. They are so designed that an EcoRI site is formed at the 5'-end and a BamHI site is formed at the 3'-end of the DNA fragments amplified by these two primers.

When PCR reaction was carried out using the genomic DNA of C. boidinii as a template and the above two oligonucleotides as a primer, an amplified DNA fragment of about 1 kb was observed. The amplified fragment was recovered and a DNA fragment of about 250 bp obtained by digestion of said fragment with a restriction enzyme EcoRI was inserted into the EcoRI-digested pBluescript II SK+. Analysis of the nucleotide sequence of the inserted fragment revealed a nucleotide sequence encoding an amino acid sequence having a high homology with the amino acid sequence of the PDI of S. cerevicesae, and therefore this DNA fragment was concluded to be part of the PDI gene of C. boidinii.

(2) Southern hybridization analysis of the genomic DNA

Genomic DNA was isolated from the bacterial cells of Candida boidinii strain S2. As the method for isolating DNA, there is mentioned a method by Cryer (Cryer, D.R. et al. (1975) Meth. Cell. Biol. 12: 39-44). The genomic DNA of Candida boidinii strain S2 was cleaved with various restriction enzymes and then separated on a 0.7% agarose gel by electrophoresis. The separated DNA was transferred to and immobilized on a nylon membrane (manufactured by Amersham). A 250 bp DNA fragment containing the above-mentioned PDI gene was labelled with ^{32}P using the Random Primer kit (manufactured by Amersham).

The labelled DNA fragment was added to a 5 x SSC - 1% SDS - 1 x Denhardt solution to prepare a hybridization solution. This hybridization solution was added to the DNA-immobilized nylon membrane and encapsulated in a plastic bag. After the encapsulated plastic bag was incubated at 65°C for 16 hours, the nylon membrane was removed from the plastic bag and washed in a 2 x SSC - 0.1% SDS solution at room temperature. Subsequently the nylon membrane was incubated in a 0.2 x SSC - 0.1% SDS solution and after the solution was replaced with a new one incubation at 65°C for 30 minutes was repeated. After the membrane was washed in a 2 x SSC, it was air-dried and subjected to autoradiography. As the smallest DNA fragment hybridizing to the above-mentioned 250 bp probe, an XbaI fragment of about 6.2 kb was found as shown in Fig. 2.

(3) Cloning of the PDI gene by colony hybridization

The genomic DNA of Candida boidinii strain S2 was completely digested with a restriction enzyme XbaI and fractionated on a 0.7% agarose gel electrophoresis. The agarose at around 6.2 kb was excised and the DNA fragment was recovered using a DNA cell (manufactured by Daiichi Kagaku). The recovered DNA was inserted into the XbaI-digested pBluescript II SK+, and Escherichia coli strain JM109 was transformed to prepare the genomic library of Candida boidinii strain S2.

The library was screened by colony hybridization using the above-mentioned 250 bp DNA fragment as a probe to obtain positive clones. The hybridization conditions were the same as that of the above-mentioned Southern hybridization. Plasmid was recovered from the positive clones to create a restriction enzyme map of the inserted DNA fragments. The restriction enzyme map so created is shown in Fig. 1. Subcloning was carried out based on the restriction enzyme map and the DNA fragment containing the PDI gene was limited to about 2 kb (the left hand side in Fig. 1) spanning from XbaI to SalI.

(4) Determination of the nucleotide sequence

The nucleotide sequence of the above DNA fragment of about 2 kb spanning from XbaI to SalI was determined. The DNA fragment was cloned into phage M13 in the both directions to prepare each of the double stranded DNA's (RF). These double stranded DNA's were allowed to react with Escherichia coli exonuclease III to prepare a double stranded DNA in which deletion has been introduced in one direction. A method for making an plasmid having a one-direction deletion insertion using exonuclease III has been described in detail on pages 289-305 in "Zoku Seikagaku Jikken Kouza (Sequel to the Series of Biochemistry Experiments), Vol. 1, Idenshi Kenkyuuhou (Methods for Studying Genes) II".

Each of the double stranded DNA's in which deletion has been inserted in one direction obtained in the above method was transformed into E. coli strain JM109 to make a phage clone in which deletion has been inserted in one

direction. From each phage clone a double stranded DNA was prepared, for which the degree of deletion was investigated from the cleavage pattern by restriction enzymes, and then single stranded phage DNA's were prepared from appropriate clones. Using these single stranded phage DNA's as the template, the nucleotide sequence was determined by the dideoxy method (Sanger, F. et al. (1977) Proc. Natl. Acad. Sci. U.S.A. 74: 5463). By ligating the nucleotide sequence of each clone the nucleotide sequence of 2.0 kb spanning from the XbaI site to immediately before the SalI site in Fig. 2 was determined.

SEQ ID No. 1 shows the nucleotide sequence and the amino acid sequence of the PDI deduced from the nucleotide sequence. The PDI of C. boidinii was found to consist of 531 amino acids encoded by the nucleotide sequence from the base No. 367 through 1959 of the nucleotide sequence shown in SEQ ID No. 1, and was designated the PDI1 gene. The amino acid sequence of the PDI1 has shown a homology of 45% with the PDI derived from S. cerevicease and 22% with the human PDI. When analogous amino acids are considered, the homology was 64% with the PDI of S. cerevicease and 49% with the human PDI.

The sequence which has been conserved in the PDI of various origins as the active center of the disulfide bond exchange reaction of the PDI, i.e. Cys-Gly-His-Cys (SEQ ID No. 4), was found in two sites: the amino acid sequence from amino acids 61 to 64 and that from amino acids 408 to 411 of the amino acid sequence of SEQ ID No. 1. Furthermore, the ER retention signal sequence present in the C-terminal was Arg-Asp-Glu-Leu (SEQ ID No. 9), which was different from the PDI of S. cerevicease, His-Asp-Glu-Leu (SEQ ID No. 3), or Lys-Asp-Glu-Leu (SEQ ID No. 2) widely occurring in the PDI of mammals.

The measurement of the activity of protein disulfide isomerase can be performed by investigating the accelerating effect on reassembly of the scrambled ribonuclease A (RNase A) which was made by a method comprising reduction, denaturation and reoxidation. The degree of reassembly of ribonuclease A is quantitated using the degree of recovery of the enzymatic activity as an index (Japanese Unexamined Patent Publication (Kokai) No. 6-38771).

By measuring the PDI activity by the above-mentioned method, it was confirmed that the strain of methylotrophic yeast transformant containing the above DNA fragment had a higher protein disulfide isomerase activity than the untransformed strain of methylotrophic yeast as the control as shown in Fig. 1.

Example 2. Secretion of the desired heterologous protein

It was confirmed that the amount secreted of ARP is increased by coexpressing the PDI1 gene derived from the strain of methylotrophic yeast C. boidinii and the peroxidase gene (ARP) gene derived from a filamentous fungus Arthromyces ramosus. pNOTe1 used as an expression vector and the ARP expression vector pNOTe1ARP have been disclosed in Japanese Unexamined Patent Publication (Kokai) No. 5-344895. By exchanging the auxotrophic marker (URA3) of pNOTe1 for the LEU2 gene derived from C. boidinii, an expression vector having an auxotrophic marker different from pNOTe1 can be created. It is also possible to effect transformation by the two expression vectors mentioned above, by imparting to the strain of methylotrophic yeast auxotrophy corresponding to the markers of these two expression vectors. As a method for imparting auxotrophy to the strain of methylotrophic yeast, a known method (Sakai, Y. et al. (1991) J. Bacteriol. 173: 7458-7463) can be used.

(1) Construction of expression vectors

A 1.1 kb EcoRI DNA fragment containing the ARP gene was excised from plasmid pNOTe1ARP, and then inserted into the NotI site of pNOTe1 to create plasmid pNPO3 as shown in Fig. 3 (a).

For the purpose of expressing the PDI1 gene, an expression vector having the LEU2 gene as an auxotrophic marker and the ribosome DNA (rDNA) of C. boidinii as a recombination site was created in the procedure as set forth in Fig. 4. To begin with, pNOTe1 was cleaved with EcoRI and HindIII and then a 2.0 kb DNA fragment containing the promoter and terminator of the alcohol oxidase gene (AOD1) of C. boidinii was excised and inserted into the EcoRI-HindIII site of pUC19 to create plasmid pNOT46. A DNA fragment containing rDNA derived from C. boidinii was obtained by the PCR method and was then inserted into the HindIII site of pNOT46 to create pNOT46R. Plasmid pCLEU321 (Sakai, Y. et al. (1992) J. Bacteriol. 174: 5988-5993) containing the LEU2 gene of C. boidinii was digested with EcoRI, and a DNA fragment containing a 3.2 kb LEU2 gene, which was rendered blunt-ended. After the blunt-ended 3.2 kb DNA fragment was digested with NdeI, it was inserted into the blunt-ended pNOT46R to create pNL1.

In order to integrate the above expression vectors, a NotI site was created on both ends of the PDI gene by the PCR method. As the sense primer,

5'-ATAAGAATGCGGCCGCAAAATGAAGTTAACTAATTTCAAA-3' (SEQ ID No. 10),

and as the antisense primer,

5'-ATAAGAATGCGGCCGCTTATAATTCATCACGAACATCA-3' (SEQ ID No. 11)

were synthesized. At the 5'-end of these two oligonucleotides there is a sequence recognized by NotI so that a NotI site may be created immediately before the initiation codon and immediately after the termination codon of the PDI1 gene in a DNA fragment amplified using these primers. Using the genomic DNA of C. boidinii as a template and the above two primers as a primer, PCR reaction was carried out, and the amplified 1.6 kb DNA fragment was digested with NotI, which was inserted into the NotI site of plasmid pBluescript II SK+ to create pSKPD. A 1.6 kb DNA fragment obtained by digesting pSKPD with NotI was inserted into the NotI site of the above pNL1 to create pNRPD as shown in Fig. 3 (b).

(2) Creation of a transformed yeast

Using the two expression vectors mentioned above in (1), a transformant of the strain of methylotrophic yeast C. boidinii was created. The bacterial strain used as the host is C. boidinii BUL (ura3, leu2) wherein the LEU2 gene of C. boidinii strain TK62 (ura3) (disclosed in Japanese Unexamined Patent Publication (Kokai) No. 5-344895) has been destructed. The LEU2 gene of C. boidinii has been disclosed by Sakai et al. (Sakai, Y. et al. (1992) J. Bacteriol. 174: 5988-5993). The method of transformation of C. boidinii has been disclosed in Japanese Unexamined Patent Publication (Kokai) No. 5-344895.

To begin with, a transformant of the ARP gene was created. After the ARP expression vector pNPO3 was linearized by digestion with BamHI, C. boidinii strain BUL (ura3, leu2) was transformed, and transformants were selected using Ura3+. As shown in Fig. 5 A), B), in one of the transformants selected, the BPO17 strain, the entire region of pNPO3 containing the ARP gene has been integrated into the ura3 site by homologous recombination of the ura3 site on the chromosomal DNA of the host yeast BUL strain and the URA3 site in the expression vector pNPO3. This was confirmed by the fact, as shown in Fig. 5 C), that in Southern hybridization carried out using as a probe a 3.3 kb BamHI-SalI DNA fragment containing the entire region of the URA3 gene after the genomic DNA of the host BUL strain and the transformant BPO17 strain were digested with BglII, a 5.5 kb hybridizing band in the BUL strain and a 14.4 kb hybridizing band in the BPO17 strain were observed.

Next, using as the host C. boidinii strain BPO17 (leu2) which was transformed with the above-mentioned ARP gene, a transformant of the PDI1 gene was created. After the PDI1 expression vector pNRPD was linearized by digestion with ApaI, the BPO17 (leu2) strain was transformed and the BPP1 strain was obtained after selection with Leu+. In the BPP1 strain, as shown in Fig. 6, A), B), the entire region of pNRPD containing the PDI1 gene has been integrated into the rDNA site by homologous recombination of the rDNA site on the chromosomal DNA of the host yeast BPO17 strain and the rDNA site in the expression vector pNPO3. Integration of the PDI1 gene into the chromosomal DNA was confirmed by the fact, as shown in Fig. 6 C), that in Southern hybridization carried out using as a probe a 1.6 kb DNA fragment containing the PDI1 gene obtained by digestion of pSKPD with NotI after the genomic DNA of the BUL strain, the host BPO17 strain and the transformant BPP1 strain were digested with HindIII, a band derived from the region containing a 12.6 kb intrinsic PDI1 gene from the BUL strain and the BPO17 strain and a 6.1 kb band derived from the expression vector pNRPD in addition to the above 12.6 kb band from the BPP1 strain were observed.

(3) Analysis of transformants

mRNA was extracted from the BPO17 strain transformed with the ARP gene, the BPP1 strain transformed with the ARP gene and the PDI1 gene, and the BUL strain used as the host, and the amount expressed of the PDI1 gene was investigated by Northern hybridization. From the bacterial cells obtained from the above three strains cultured at 30°C for 48 hours in the YM medium having methanol as the sole carbon source (Sakai, Y. et al. (1991) J. Gen. Microbiol. 123, 365-398), total RNA was extracted by ISOGEN (manufactured by Nihon Gene K.K.) and purified using BIOMAG mRNA purification kit (manufactured by PerSeptive Diagnostics). The purified mRNA was subjected to a 1.1% agarose gel electrophoresis (containing 20 mM MOPS buffer, 1 mM EDTA, 2.2 M formamide), and then blotted onto the nylon membrane. In the same condition as the Southern hybridization described in Example 1, hybridization was carried out. The probe used in the hybridization was 1.6 kb NotI DNA fragment derived from the above-mentioned pSKPD.

As shown in Fig. 7, strong expression of the PDI1 gene was observed in the BPP1 strain which was transformed with the PDI1 gene and weak expression of possibly the intrinsic PDI1 gene was observed in the BUL strain and the

BPO17 strain.

The above three bacterial strains were cultured in the YM medium containing methanol as the sole carbon source at 30°C for 48 hours, and then the PDI activity in the bacterial cells was measured. The harvested cells were suspended in 50 mM potassium phosphate buffer, pH 7.5, and transferred into a 2 ml Eppendorf tube, to which was added an equal volume of zirconium beads (0.5 mm in diameter). A procedure of vigorous stirring of the tube for 30 seconds using the Beads Beader (Model 3110BX, Biospec Products) followed by cooling on ice for 30 seconds was repeated for six times. The disrupted cells were centrifuged at 4°C at 16,000 x g for 5 minutes and then the enzymatic activity of the supernatant was measured.

The measurement of the PDI activity was carried out in accordance with the method of Hillson et al. (Hillson, D. A. et al. (1984) Methods Enzymol. 107: 281-294). Thus, one ml of the final reaction mixture contains 50 mM potassium phosphate buffer, pH 7.5, 500 µg of scrambled RNase, and 0.01 mM of dithiothreitol. After the reaction mixture was incubated for 10 minutes, 10 µl was sampled out, to which 3 ml of the TKM buffer (50 mM Tris-HCl, pH 7.5, 25 mM KCl, 5 mM MgCl) containing 0.25 mg yeast RNA was added and then RNase activity was determined by measuring absorbance at 260 nm in a UV cuvette at 30°C for 2 minutes. One unit of the enzymatic activity was defined as the amount of enzyme which increases the absorbance at 260 nm per one minute.

As shown in Table 1, the PDI activity in the bacterial cells was higher in the BPP1 strain transformed with the PDI gene than the BPO17 strain transformed with the ARP gene alone or the BUL strain used as the host by a factor of 9 or more.

Table 1

Enzyme \ Strain	BUL	BPO17	BPP1
PDI	<0.1*	<0.1*	0.896

* The levels of BUL and BPO17 were below the detection

(ml)

(4) Secretory expression of the ARP

The BPO17 strain transformed with the ARP gene, the BPP1 strain transformed with both of the ARP gene and the PDI1 gene, and the BUL strain used as the host were cultivated in the YM medium containing methanol as the sole carbon source, and the ARP activity in the culture liquid was compared. As shown in Fig. 8, the ARP activity in the culture liquid of the BPP1 strain coexpressed with the ARP gene and the PDI1 gene reached a maximum of 0.024 U/ml at 84 hours after cultivation. In the BPO17 strain in which the ARP gene only was expressed, the ARP activity in the culture liquid reached a maximum of 0.002 U/ml at 84 hours after cultivation, while no ARP activity was observed in the culture liquid of the BUL strain used as the host. The result revealed that by coexpressing the PDI1 gene and the ARP gene the amount secreted of ARP increased by about 10 fold.

The present invention made it possible to obtain the PDI gene of the strain of methylotrophic yeast and to obtain the PDI enzyme by expressing said gene in large quantities using said strain of methylotrophic yeast. Furthermore, by coexpressing said gene with the gene of the desired secretory protein in the strain of methylotrophic yeast it became possible to drastically increase the amount produced of the desired protein.

Note: this invention extends to mutants or modified forms of the nucleotide sequence and protein sequence of SEQ. ID. 1, provided that these are associated with the protein disulfide isomerase activity, which at the least includes the ability to catalyse formation/exchange of protein disulfide bonds and location in the ER.

Likewise, the use of vectors containing such mutants (modified sequences to transform hosts so as to increase their PDI activity in particular for expressing a recombinant (heterologous) protein.

SEQUENCE LISTING

SEQ ID No: 1

Sequence length: 2030

Sequence type: nucleic acid

Strandedness: double

Topology: linear

Molecule type: genomic DNA

Hypothetical: No

Antisense: No

Original source:

Organism: Candida boidinii

Strain: S2

Sequence description:

```
AGAGCGCTCT CGACTCACTC ATTATTCATC CAGTATCTGC TCCAAGGTTG TGAACAATTT   60
CACTCACTTG CCTTGCTTTA CCATCTACTCA ATCGTTTCA TTTACTCCTG TATCATTCCA  120
CCATTTCATC ACTTTTTCAT ATCTAGTAACT AAATGTCTA AGCAACGATA ATCTTTCAGC  180
AGATTCGCTC TTCTTTGATT CAATTGATCCT TTCATAGAC AGATCACTGA CACGTAAATA  240
CTTACATAGA TATATATATA TATATATGTAA ATTTACTTT CGTCATTACT CAATTGATTC  300
CATTTAATAC ATTCATAGTA TAATATATTGA CTTAAAATAT ATTTACATAT ACACATAACA  360
TTTAAA ATG AAG TTA ACT AAT TTC AAA GTT ATT GCC ACA ATT CTT GCT       408
      Met Lys Leu Thr Asn Phe Lys Val Ile Ala Thr Ile Leu Ala
        1               5                  10
TGT TTA ACA GTT GTT AGA GCT GAT GAT GGT GGT GCC ATT GCA TCT CCA      456
Cys Leu Thr Val Val Arg Ala Asp Asp Gly Gly Ala Ile Ala Ser Pro
 15                  20                  25                  30
GAT TCC GCT GTT GTT AAA TTA ACT GCT GAT TCA TTC GAA TCA TTC ATG      504
Asp Ser Ala Val Val Lys Leu Thr Ala Asp Ser Phe Glu Ser Phe Met
                 35                  40                  45
AAA GAA AAT CCA TTA GTC TTA GCT GAA TTT TTT GCT CCT TGG TGT GGT      552
Lys Glu Asn Pro Leu Val Leu Ala Glu Phe Phe Ala Pro Trp Cys Gly
             50                  55                  60
CAT TGT AAA AGA TTG GGT CCT GAA TTT CAA GTT GCT GCT GAT AAA TTA      600
His Cys Lys Arg Leu Gly Pro Glu Phe Gln Val Ala Ala Asp Lys Leu
         65                  70                  75
```

```
GTT GAA AAA GAT ATT AGA TTA GCT CAA ATT GAT TGT ACC GAA GAA AAA      648
Val Glu Lys Asp Ile Arg Leu Ala Gln Ile Asp Cys Thr Glu Glu Lys
 80                  85                  90

GAT TTA TGT TCT TCT TAT GGT ATT AAA GGT TAC CCA ACT TTA AAA GTC      696
Asp Leu Cys Ser Ser Tyr Gly Ile Lys Gly Tyr Pro Thr Leu Lys Val
 95                 100                 105                 110

TTT AGA GGT TAC GAA AAT GAA CCT TCT GAT TAT GCT GGT CAA AGA ACT      744
Phe Arg Gly Tyr Glu Asn Glu Pro Ser Asp Tyr Ala Gly Gln Arg Thr
                115                 120                 125

TCA GAT TCA ATC ATT TCT TAT ATG GTT AAA CAA TCA ACC CCA CCT GTC      792
Ser Asp Ser Ile Ile Ser Tyr Met Val Lys Gln Ser Thr Pro Pro Val
            130                 135                 140

TCC ATC GTT GAT GAT CTC TCA GAT ATC GAA GAT ACA ATT AAA GAA TCA      840
Ser Ile Val Asp Asp Leu Ser Asp Ile Glu Asp Thr Ile Lys Glu Ser
        145                 150                 155

AAT GAT CCT GTC TTT ATT CAA GTC TTA CCA AAA GGT TCT AAA TCT GTT      888
Asn Asp Pro Val Phe Ile Gln Val Leu Pro Lys Gly Ser Lys Ser Val
    160                 165                 170

GAA GCC GGT AAC TCA ACT TTC TTT GAA ATC GCT AAT GGT TTA AGA GAT      936
Glu Ala Gly Asn Ser Thr Phe Phe Glu Ile Ala Asn Gly Leu Arg Asp
175                 180                 185                 190

AAC TAC TCT TTT ATT TCA ACA ACA AGT ACT GAA TTC TCT TCA AAA TAC      984
Asn Tyr Ser Phe Ile Ser Thr Thr Ser Thr Glu Phe Ser Ser Lys Tyr
                195                 200                 205

TTG AAA GGT ATT AAA AAA TCA GAT ACT CCA TCT TAT ATT CTC TTT AGA     1032
Leu Lys Gly Ile Lys Lys Ser Asp Thr Pro Ser Tyr Ile Leu Phe Arg
            210                 215                 220

CCA AAT GAA GAA TTG TCT GAT GCT TCA ATC TAT AAA TTT GAT GAA ATT     1080
Pro Asn Glu Glu Leu Ser Asp Ala Ser Ile Tyr Lys Phe Asp Glu Ile
        225                 230                 235

GAT GAT ACT CAT TTA ATC GAA TTC TTA AAC GTT GAA TCA AAA CCT TTA     1128
Asp Asp Thr His Leu Ile Glu Phe Leu Asn Val Glu Ser Lys Pro Leu
    240                 245                 250

TTC GGT GAA ATG GAT GCT GGT TCT TCT TTC AAT TCT TAT GAA ATG AAA     1176
Phe Gly Glu Met Asp Ala Gly Ser Ser Phe Asn Ser Tyr Glu Met Lys
255                 260                 265                 270
```

```
TTA CCA GTT GCT TAT TAT TTC TAT AAT GAA ATC TCT GAA AAA GAT GCC  1224
Leu Pro Val Ala Tyr Tyr Phe Tyr Asn Glu Ile Ser Glu Lys Asp Ala
                275                 280                 285

GTC TCT GAT GCC ATC AGT AAA TTA GCT AAA ACT CAT AGA GGT AAA GTT  1272
Val Ser Asp Ala Ile Ser Lys Leu Ala Lys Thr His Arg Gly Lys Val
            290                 295                 300

AAT TTC GTT GGT TTA GAC GCT TCT AAA TAT GGT TTA CAC GCT AAG AAT  1320
Asn Phe Val Gly Leu Asp Ala Ser Lys Tyr Gly Leu His Ala Lys Asn
        305                 310                 315

ATT AAC ATG AAG GAA GAA TTC CCT CTT TTC GCT ATT CAC GAT TTA GCA  1368
Ile Asn Met Lys Glu Glu Phe Pro Leu Phe Ala Ile His Asp Leu Ala
    320                 325                 330

ACT GAA TTA AAA TAC GGT ATC TCC CAA GAT AAA CCA TTA GAT AAT AAA  1416
Thr Glu Leu Lys Tyr Gly Ile Ser Gln Asp Lys Pro Leu Asp Asn Lys
335                 340                 345                 350

TTA ATT CCA AAA TTC GTT GAA GAT TTC GTT GCT GGT AAA TTA GAA GCA  1464
Leu Ile Pro Lys Phe Val Glu Asp Phe Val Ala Gly Lys Leu Glu Ala
            355                 360                 365

ATC ATT AAA TCA GAA CCA ATC CCA GAA GAA ACT CAA GAT TCT CCA GTT  1512
Ile Ile Lys Ser Glu Pro Ile Pro Glu Glu Thr Gln Asp Ser Pro Val
        370                 375                 380

CAT TTA GTC GGT AAA GAA CAT GAA AAA ATT ACC TCT GAT AAA GAT GTC  1560
His Leu Val Gly Lys Glu His Glu Lys Ile Thr Ser Asp Lys Asp Val
    385                 390                 395

GTC TTA GTT AAA TAT TAC TAT GCT CCA TGG TGT GGT CAC TGT AAA AAA  1608
Val Leu Val Lys Tyr Tyr Tyr Ala Pro Trp Cys Gly His Cys Lys Lys
400                 405                 410

GCT CCA GTC TTT GAA GAA TTA GCT GCT GTT TAT GAA TCA GTT GCT CCA  1656
Ala Pro Val Phe Glu Glu Leu Ala Ala Val Tyr Glu Ser Val Ala Pro
415                 420                 425                 430

GGT AAA GTC TTA TTA GCT GAT TTA GAT GCT AAT CAT ACT GAA AAT GAT  1704
Gly Lys Val Leu Leu Ala Asp Leu Asp Ala Asn His Thr Glu Asn Asp
            435                 440                 445

GGT GTT CAC ATT GAA GGT TAC CCA ACT ATC GTC TTA TAC CCA GCC GAT  1752
Gly Val His Ile Glu Gly Tyr Pro Thr Ile Val Leu Tyr Pro Ala Asp
        450                 455                 460
```

GGT TCA GAA CCA GTT GTT TAC GAA GCT AAC AGA TCT TTT GAA TCT TTC 1800
Gly Ser Glu Pro Val Val Tyr Glu Gly Asn Arg Ser Phe Glu Ser Phe
465 470 475

TCC GAT TTC ATT AAA GAA AAA GGT TCA TCA GGT GTT GAT GCT AAT GCA 1848
Ser Asp Phe Ile Lys Glu Lys Gly Ser Ser Gly Val Asp Ala Asn Ala
480 485 490

TTA AAA GAA CCT TAC CCA GAA GAA GGT ACT GAA GGT GCT CCA GTT GAT 1896
Leu Lys Glu Pro Tyr Pro Glu Glu Gly Thr Glu Gly Ala Pro Val Asp
495 500 505 510

CCA GAA TCA GTT GGT GAT GCT GAA AAA GAA GAT GAT TCT GCT GCT GAT 1944
Pro Glu Ser Val Gly Asp Ala Glu Lys Glu Asp Asp Ser Ala Ala Asp
515 520 525

GTT CGT GAT GAA TTA TAAACAAGTA GAATTAATTA TAAATTGATT AAATAGTCTT 1999
Val Arg Asp Glu Leu
530 531

CTAAAAATTA AATTTAAAAT AATAAAAAAA A 2030

SEQ ID No: 2

Sequence length: 4

Sequence type: amino acid

Topology: linear

Molecule type: peptide

Sequence description:

Lys Asp Glu Leu

SEQ ID No: 3

Sequence length: 4

Sequence type: amino acid

Topology: linear

Molecule type: peptide

Sequence description:

His Asp Glu Leu

SEQ ID No: 4

Sequence length: 4

Sequence type: amino acid

Topology: linear

Molecule type: peptide

Sequence description:

Cys Gly His Cys

SEQ ID No: 5

Sequence length: 7

Sequence type: amino acid

Topology: linear

Molecule type: peptide

Sequence description:

Pro Trp Cys Gly His Cys Lys
5

SEQ ID No: 6

Sequence length: 7

Sequence type: amino acid

Topology: linear

Molecule type: peptide

Sequence description:

Tyr Ala Pro Trp Cys Gly His
5

SEQ ID No: 7

Sequence length: 29

Sequence type: nucleic acid

Topology: linear

Molecule type:

Sequence description:

CCGGAATTCC CWTGGTGTYGG WCAYTGYAA

SEQ ID No: 8

Sequence length: 28

Sequence type: nucleic acid

Topology: linear

Molecule type:

Sequence description:

CGCGGATCCT GWCCRCACCA WGGDGCRT

SEQ ID No: 9

Sequence length: 4

Sequence type: amino acid

Topology: linear

Molecule type: peptide

Sequence description:

Arg Asp Glu Leu

SEQ ID No: 10

Sequence length: 40

Sequence type: nucleic acid

Topology: linear

Molecule type: chemical synthetic DNA

Sequence description:

ATAAGAATGC GGCCGCAAAA TGAAGTTAAC TAATTTCAAA 40

SEQ ID No: 11

Sequence length: 38

Sequence type: nucleic acid

Topology: linear

Molecule type: chemical synthetic DNA

Sequence description:

ATAAGAATGC GGCCGCTTAT AATTCATCAC GAACATCA 38

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Suntory Limited
(B) STREET: 1-40, Dojimahama 2-chome, Kita-ku, Osaka-shi
(C) CITY: Osaka
(E) COUNTRY: Japan
(F) POSTAL CODE (ZIP): Osaka 530

(ii) TITLE OF INVENTION: Protein disulfide isomerase gene derived from strain of methylotrophic yeast

(iii) NUMBER OF SEQUENCES: 11

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: EP 97306871.1

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: JP 8-234287
(B) FILING DATE: 04-SEP-1996

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2030 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Candida boidinii
(B) STRAIN: S2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AGAGCCTCT CCACTCACTC ATTATTCATC CAGTATCTCC TCCAAGGTTG TGAACAATTT 60
CACTCACTTG CCTTGCTTTA CCATCTACTCA ATCGTTTCA TTTACTCCTG TATCATTCCA 120
CCATTTCATC ACTTTTTCAT ATCTAGTAACT AAATGTCTA AGCAACGATA ATCTTTCAGC 180
AGATTCCCTC TTCTTTGATT CAATTGATCCT TTCATAGAC AGATCACTGA CACGTAAATA 240
CTTACATAGA TATATATATA TATATATGTAA ATTTACTTT CGTCATTACT CAATTGATTC 300
CATTTAATAC ATTCATAGTA TAATATATTGA CTTAAATAT ATTTACATAT ACACATAACA 350
TTTAAA ATG AAG TTA ACT AAT TTC AAA GTT ATT GCC ACA ATT CTT GCT 403
Met Lys Leu Thr Asn Phe Lys Val Ile Ala Thr Ile Leu Ala
1 5 10
TGT TTA ACA GTT GTT AGA GCT GAT GAT GGT GGT GCC ATT GCA TCT CCA 451
Cys Leu Thr Val Val Arg Ala Asp Asp Gly Gly Ala Ile Ala Ser Pro
15 20 25 30
GAT TCC GCT GTT GTT AAA TTA ACT GCT GAT TCA TTC GAA TCA TTC ATG 504
Asp Ser Ala Val Val Lys Leu Thr Ala Asp Ser Phe Glu Ser Phe Met
35 40 45
AAA GAA AAT CCA TTA GTC TTA GCT GAA TTT TTT GCT CCT TGG TGT GGT 552
Lys Glu Asn Pro Leu Val Leu Ala Glu Phe Phe Ala Pro Trp Cys Gly
50 55 60
CAT TGT AAA AGA TTG GGT CCT GAA TTT CAA GTT GCT GCT GAT AAA TTA 600
His Cys Lys Arg Leu Gly Pro Glu Phe Gln Val Ala Ala Asp Lys Leu
65 70 75
GTT GAA AAA GAT ATT AGA TTA GCT CAA ATT GAT TGT ACC GAA GAA AAA 648
Val Glu Lys Asp Ile Arg Leu Ala Gln Ile Asp Cys Thr Glu Glu Lys
80 85 90
GAT TTA TGT TCT TCT TAT GGT ATT AAA GGT TAC CCA ACT TTA AAA GTC 696
Asp Leu Cys Ser Ser Tyr Gly Ile Lys Gly Tyr Pro Thr Leu Lys Val
95 100 105 110
TTT AGA GGT TAC GAA AAT GAA CCT TCT GAT TAT GCT GGT CAA AGA ACT 744
Phe Arg Gly Tyr Glu Asn Glu Pro Ser Asp Tyr Ala Gly Gln Arg Thr
115 120 125
TCA GAT TCA ATC ATT TCT TAT ATG GTT AAA CAA TCA ACC CCA CCT GTC 792
Ser Asp Ser Ile Ile Ser Tyr Met Val Lys Gln Ser Thr Pro Pro Val
130 135 140
TCC ATC GTT GAT GAT CTC TCA GAT ATC GAA GAT ACA ATT AAA GAA TCA 840
Ser Ile Val Asp Asp Leu Ser Asp Ile Glu Asp Thr Ile Lys Glu Ser
145 150 155
AAT GAT CCT GTC TTT ATT CAA GTC TTA CCA AAA CCT GGT TCT AAA TCT GTT 888
Asn Asp Pro Val Phe Ile Gln Val Leu Pro Lys Pro Gly Ser Lys Ser Val
160 165 170

```
GAA GCC GGT AAC TCA ACT TTC TTT GAA ATC GCT AAT GGT TTA AGA GAT      936
Glu Ala Gly Asn Ser Thr Phe Phe Glu Ile Ala Asn Gly Leu Arg Asp
175                 180                 185                 190

AAC TAC TCT TTT ATT TCA ACA ACA AGT ACT GAA TTC TCT TCA AAA TAC      984
Asn Tyr Ser Phe Ile Ser Thr Thr Ser Thr Glu Phe Ser Ser Lys Tyr
                195                 200                 205

TTG AAA GGT ATT AAA AAA TCA GAT ACT CCA TCT TAT ATT CTC TTT AGA     1032
Leu Lys Gly Ile Lys Lys Ser Asp Thr Pro Ser Tyr Ile Leu Phe Arg
            210                 215                 220

CCA AAT GAA GAG TTG TCT GAT GCT TCA ATC TAT AAA TTT GAT GAA ATT     1080
Pro Asn Glu Glu Leu Ser Asp Ala Ser Ile Tyr Lys Phe Asp Glu Ile
        225                 230                 235

GAT GAT ACT CAT TTA ATC GAA TTC TTA AAC GTT GAA TCA AAA CCT TTA     1128
Asp Asp Thr His Leu Ile Glu Phe Leu Asn Val Glu Ser Lys Pro Leu
    240                 245                 250

TTC GGT GAA ATG GAT GGT TCT TCT TTC CAA TCT TAT ATG GAA ATG AAA     1176
Phe Gly Glu Met Asp Gly Ser Ser Phe Gln Ser Tyr Met Glu Met Lys
255                 260                 265                 270

TTA CCA GTT GCT TAT TAT TTC TAT AAT GAA ATC TCT GAA AAA GAT GCC     1224
Leu Pro Val Ala Tyr Tyr Phe Tyr Asn Glu Ile Ser Glu Lys Asp Ala
            275                 280                 285

GTC TCT CAT GCC ATC AGT AAA TTA GCT AAA ACT CAT AGA GGT AAA GTT     1272
Val Ser Asp Ala Ile Ser Lys Leu Ala Lys Thr His Arg Gly Lys Val
        290                 295                 300

AAT TTC GTT GGT TTA GAC GCT TCT AAA TAT GGT TTA CAC GCT AAG AAT     1320
Asn Phe Val Gly Leu Asp Ala Ser Lys Tyr Gly Leu His Ala Lys Asn
    305                 310                 315

ATT AAC ATG AAG GAA GAA TTC CCT CTT TTC GCT ATT CAC GAT TTA GCA     1368
Ile Asn Met Lys Glu Glu Phe Pro Leu Phe Ala Ile His Asp Leu Ala
320                 325                 330

ACT GAA TTA AAA TAC GGT ATC TCC CAA GAT AAA CCA TTA GAT AAT AAA     1416
Thr Glu Leu Lys Tyr Gly Ile Ser Gln Asp Lys Pro Leu Asp Asn Lys
335                 340                 345                 350

TTA ATT CCA AAA TTC GTT GAA GAT TTC GTT GCT GGT AAA TTA GAA GCA     1464
Leu Ile Pro Lys Phe Val Glu Asp Phe Val Ala Gly Lys Leu Glu Ala
            355                 360                 365

ATC ATT AAA TCA GAA CCA ATC CCA GAA ACT CAA GAT TCT CCA GTT TAC     1512
Ile Ile Lys Ser Glu Pro Ile Pro Glu Thr Gln Asp Ser Pro Val Tyr
        370                 375                 380

CAT TTA GTC GGT AAA GAA CAT GAT AAA ATT ATC ACC TCT GAT AAA GAT     1560
His Leu Val Gly Lys Glu His Asp Lys Ile Ile Thr Ser Asp Lys Asp
    385                 390                 395

GTC TTA GTT AAA TAT TAC GCT CCA TGG TGT GGT CAC TGT AAA AAA TTA     1608
Val Leu Val Lys Tyr Tyr Ala Pro Trp Cys Gly His Cys Lys Lys Leu
400                 405                 410
```

```
GCT CCA GTC TTT GAA GAA TTA GCT GCT GTT TAT GAA TCA GTT GCT CCA   1656
Ala Pro Val Phe Glu Glu Leu Ala Ala Val Tyr Glu Ser Val Ala Pro
415                 420                 425                 430

GGT AAA GTC TTA TTA GCT GAT TTA GAT CAT ACT GAA AAT GAT GTC ACC   1704
Gly Lys Val Leu Leu Ala Asp Leu Asp His Thr Glu Asn Asp Val Thr
                435                 440                 445

GGT GTT CAC ATT GAA GGT TAC CCA ACT ATC GTC TTA TAC CCA GCC GAT   1752
Gly Val His Ile Glu Gly Tyr Pro Thr Ile Val Leu Tyr Pro Ala Asp
            450                 455                 460

GGT TCA GAA CCA GTT GTT TAC GAA GGT AAC AGA TCT TTT GAA TCT TTC   1800
Gly Ser Glu Pro Val Val Tyr Glu Gly Asn Arg Ser Phe Glu Ser Phe
        465                 470                 475

TCC GAT TTC ATT AAA GAA AAA GGT TCA TCA GGT GTT GAT GCT AAT GCA   1848
Ser Asp Phe Ile Lys Glu Lys Gly Ser Ser Gly Val Asp Ala Asn Ala
    480                 485                 490

TTA AAA GAA CCT TAC CCA GAA GAA GGT ACT GAA GGT GCT CCA GTT GAT   1896
Leu Lys Glu Pro Tyr Pro Glu Glu Gly Thr Glu Gly Ala Pro Val Asp
495                 500                 505                 510

CCA GAA TCA GTT GGT GAT GCT GAA AAA GAA GAT GAT TCT GCT GCT GAT   1944
Pro Glu Ser Val Gly Asp Ala Glu Lys Glu Asp Asp Ser Ala Ala Asp
                515                 520                 525

GTT CGT GAT GAA TTA TAAACAACTA CAATTAATTA TAAATTGATT AAATAGTCTT   1999
Val Arg Asp Glu Leu
            530 531

CTAAAAATTA AATTTAAAAT AATAAAAAAA A                               2030
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Lys Asp Glu Leu

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
His Asp Gln Leu

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
Cys Gly His Cys

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
Pro Trp Cys Gly His Cys Lys

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Tyr Ala Pro Trp Cys Gly His
5

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CCGGAATTCC CWTGGTGTYGG WCAYTGYAA 29

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CGCGGATCCT GWCCRCACCA WGGDGCRT 28

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Arg Asp Glu Leu

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid: chemical synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

ATAAGAATGC GGCCGCAAAA TGAAGTTAAC TAATTTCAAA 40

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid: chemical synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATAAGAATGC GGCCGCTTAT AATTCATCAC GAACATCA 38

Claims

1. A protein derived from a strain of methylotrophic yeast which has a protein disulfide isomerase activity, having the amino acid sequence as set forth in SEQ ID No. 1, or protein in which said amino acid sequence has been modified by deletion or addition of one or a few amino acids, or substitution with other amino acid(s) and which has a protein disulfide isomerase activity.

2. A protein derived from a strain of methylotrophic yeast, which has the conserved region of the active center for protein disulfide isomerase comprising Cys-Gly-His-Cys (SEQ ID No. 4) and an endoplasmic reticulum retention signal sequence comprising Arg-Asp-Glu-Leu (SEQ ID No. 9), and which has a protein disulfide isomerase activity.

3. A gene encoding a protein according to claim 1 or 2.

4. A gene according to claim 3 represented by the nucleotide sequence as set forth in SEQ ID No. 1.

5. A vector comprising a gene according to claim 3 or 4.

6. A transformant obtained by transforming a host with a vector according to claim 5.

7. A transformant according to claim 6 wherein the host is a strain of methylotrophic yeast.

8. A transformant according to claim 7 wherein the strain of methylotrophic yeast is Candida boidinii.

9. A method for producing a protein having a protein disulfide isomerase activity, which method comprises culturing a transformant according to any one of claims 6 to 8, and then recovering the protein from the culture.

10. A method for producing a peptide or a protein encoded by a heterologous structural gene, which method comprises culturing a transformant cotransformed with a vector according to claim 5 and a vector having a heterologous structural gene and then recovering an expression product of the heterologous structural gene, which is the peptide or the protein, from the culture.

11. A method for producing a peptide or a protein encoded by a heterologous structural gene, which method comprises culturing a transformant cotransformed with a vector which contains the recombinant gene of the protein disulfide isomerase of a methylotrophic yeast and a vector having a heterologous structural gene and then recovering an expression product of the heterologous structural gene, which is the peptide or the protein, from the culture.

Fig. 1

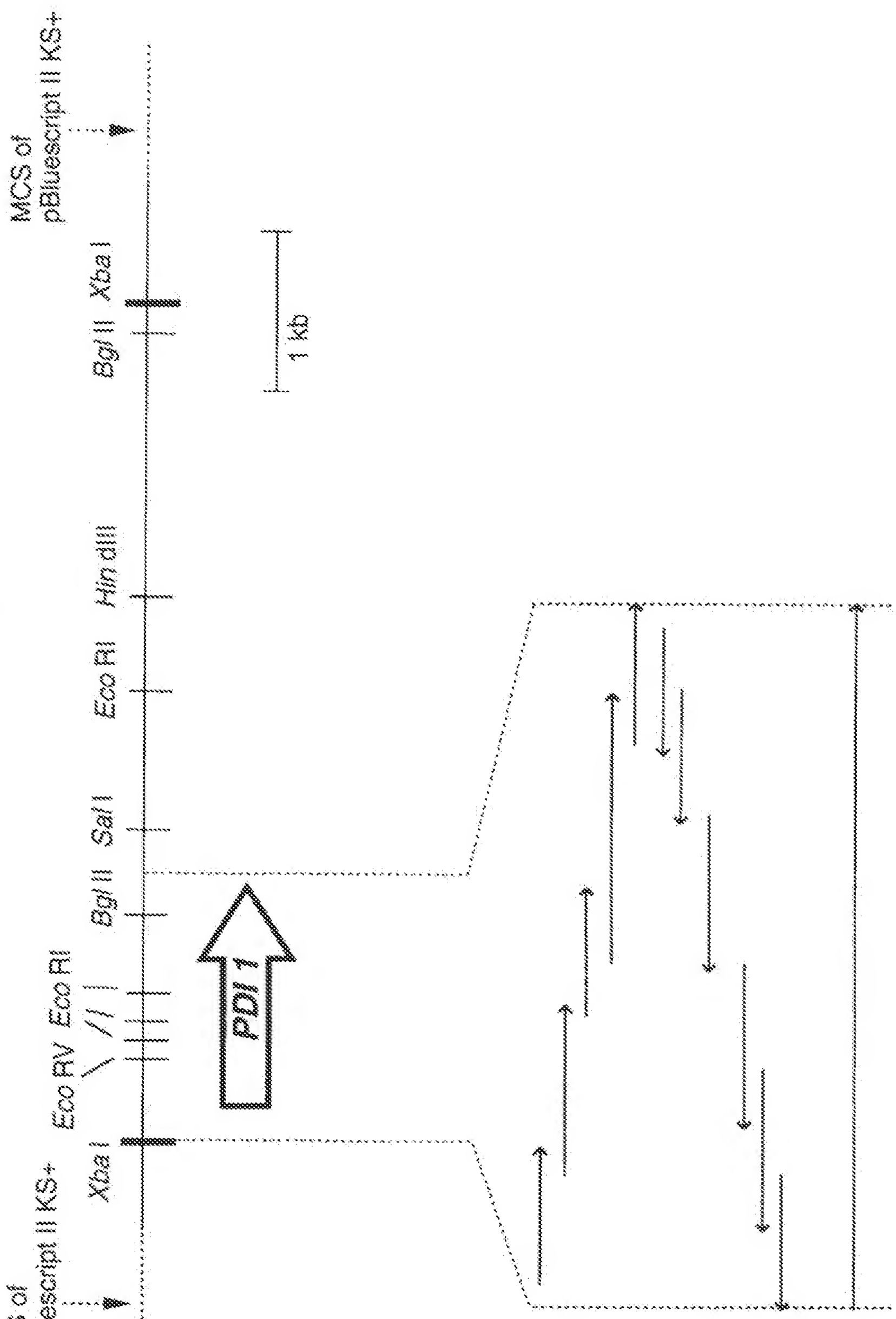

MCS of
pBluescript II KS+
Bgl II Xba I
1 kb
MCS of
pBluescript II KS+
Xba I
Eco RV Eco RI
Bgl II Sal I Eco RI Hin dIII
pD1

Fig. 2

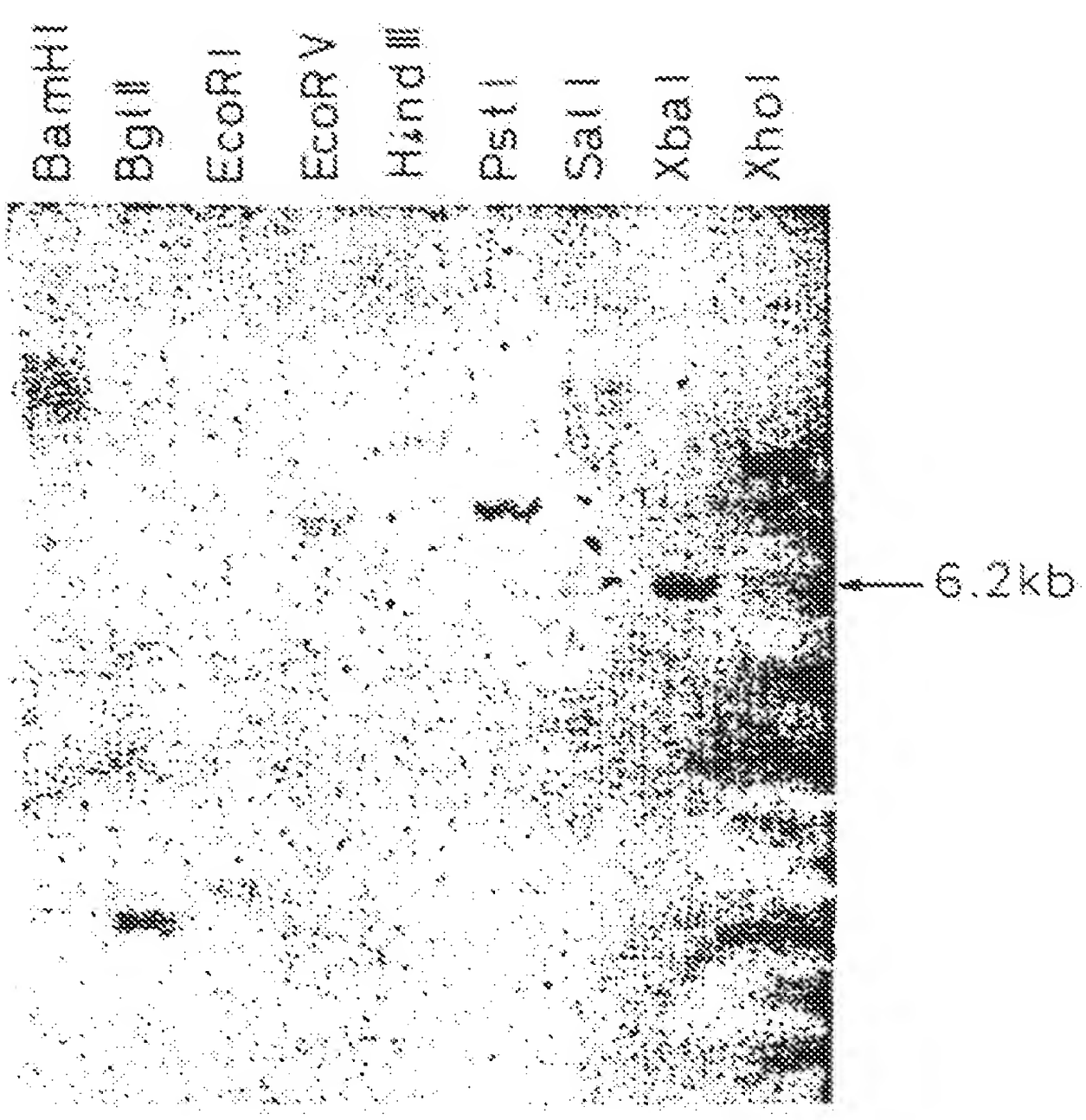
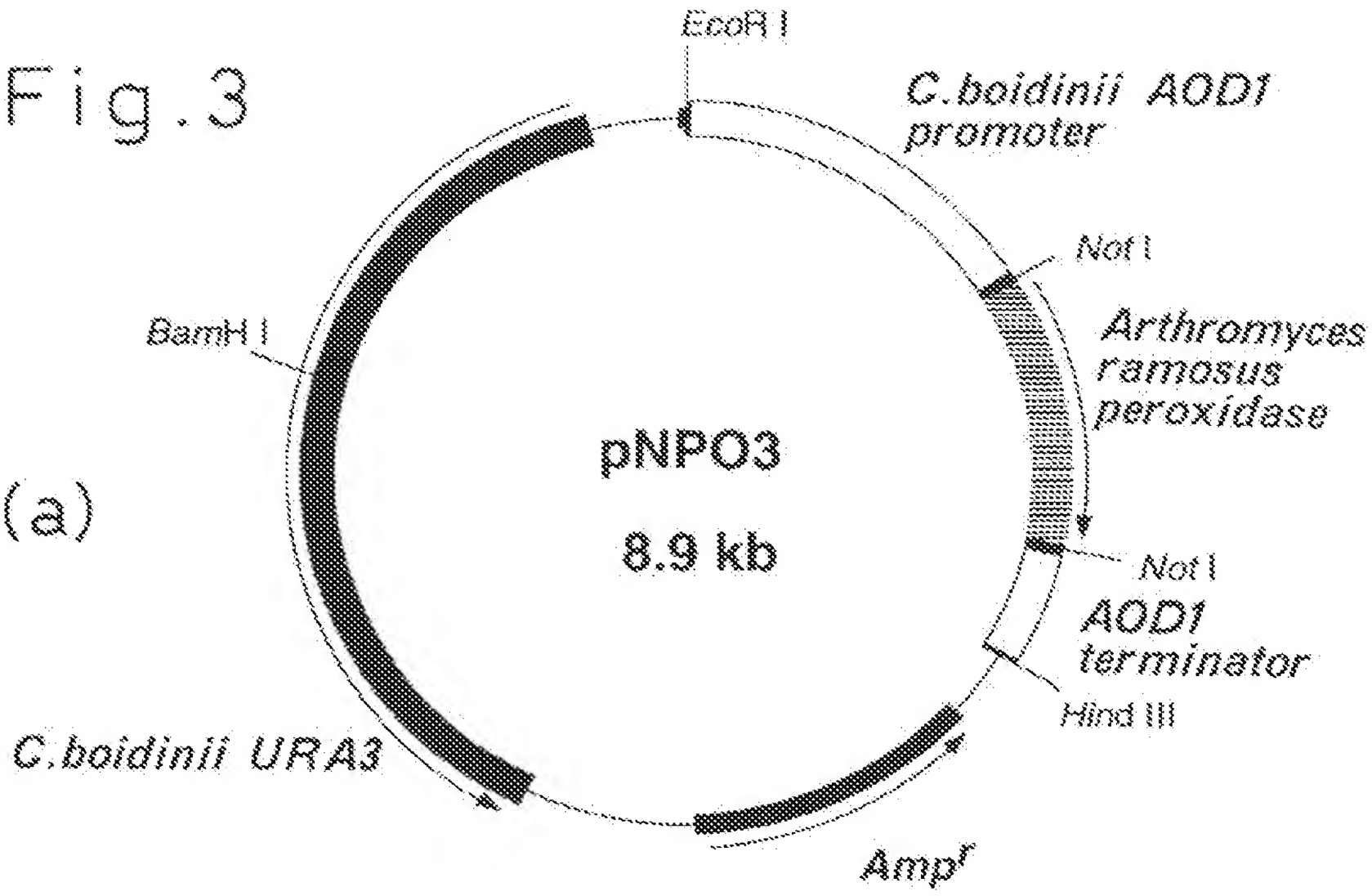


Fig.3

(a)


EcoR I
C.boidinii AOD1 promoter
Not I
Arthromyces ramosus peroxidase
BamH I
pNPO3
8.9 kb
Not I
AOD1 terminator
Hind III
C.boidinii URA3
Ampr

(b)

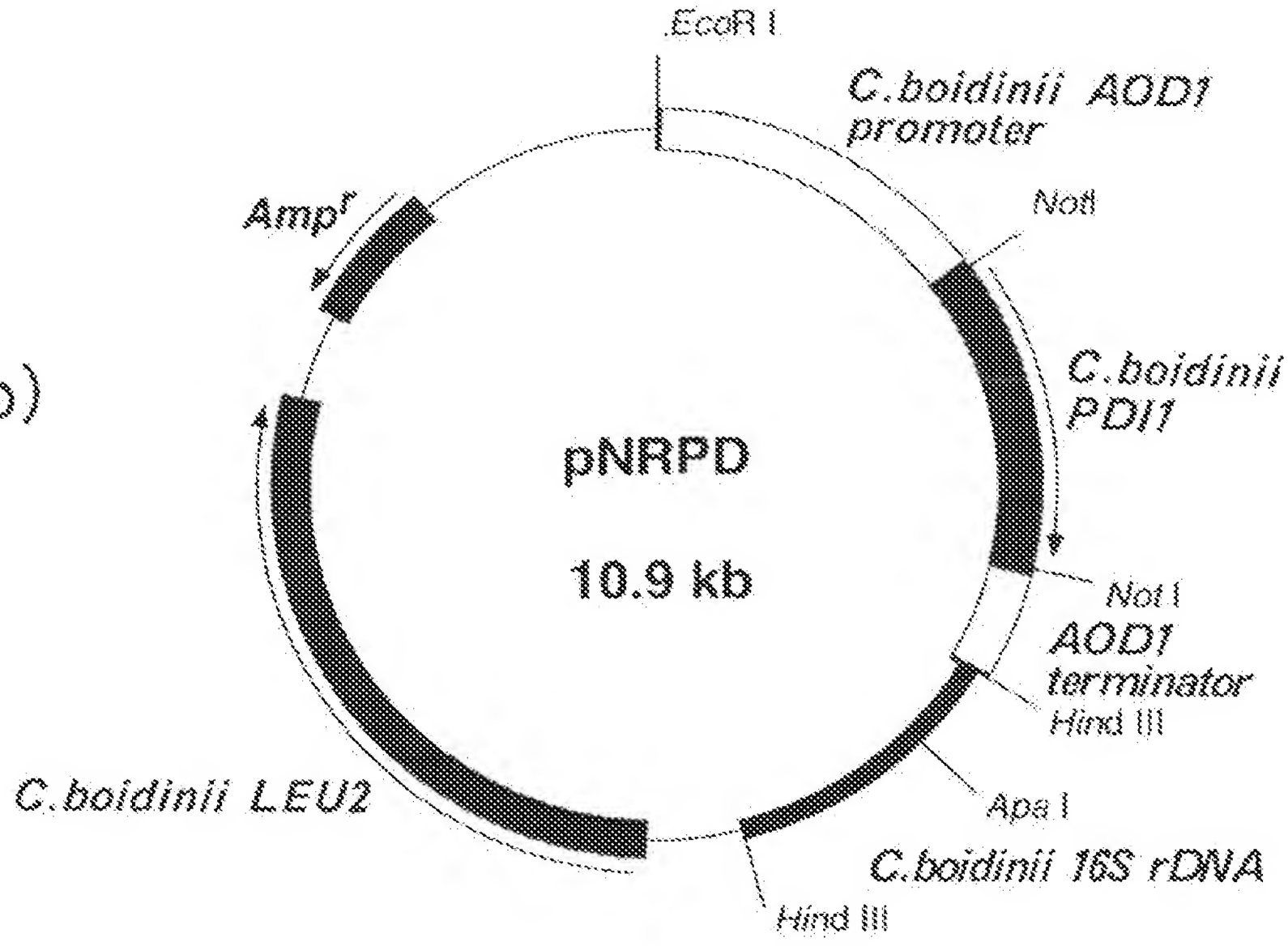

EcoR I
C.boidinii AOD1 promoter
Ampr
Not I
C.boidinii PDI1
pNRPD
10.9 kb
Not I
AOD1 terminator
Hind III
Aps I
C.boidinii LEU2
C.boidinii 16S rDNA
Hind III

Fig. 4

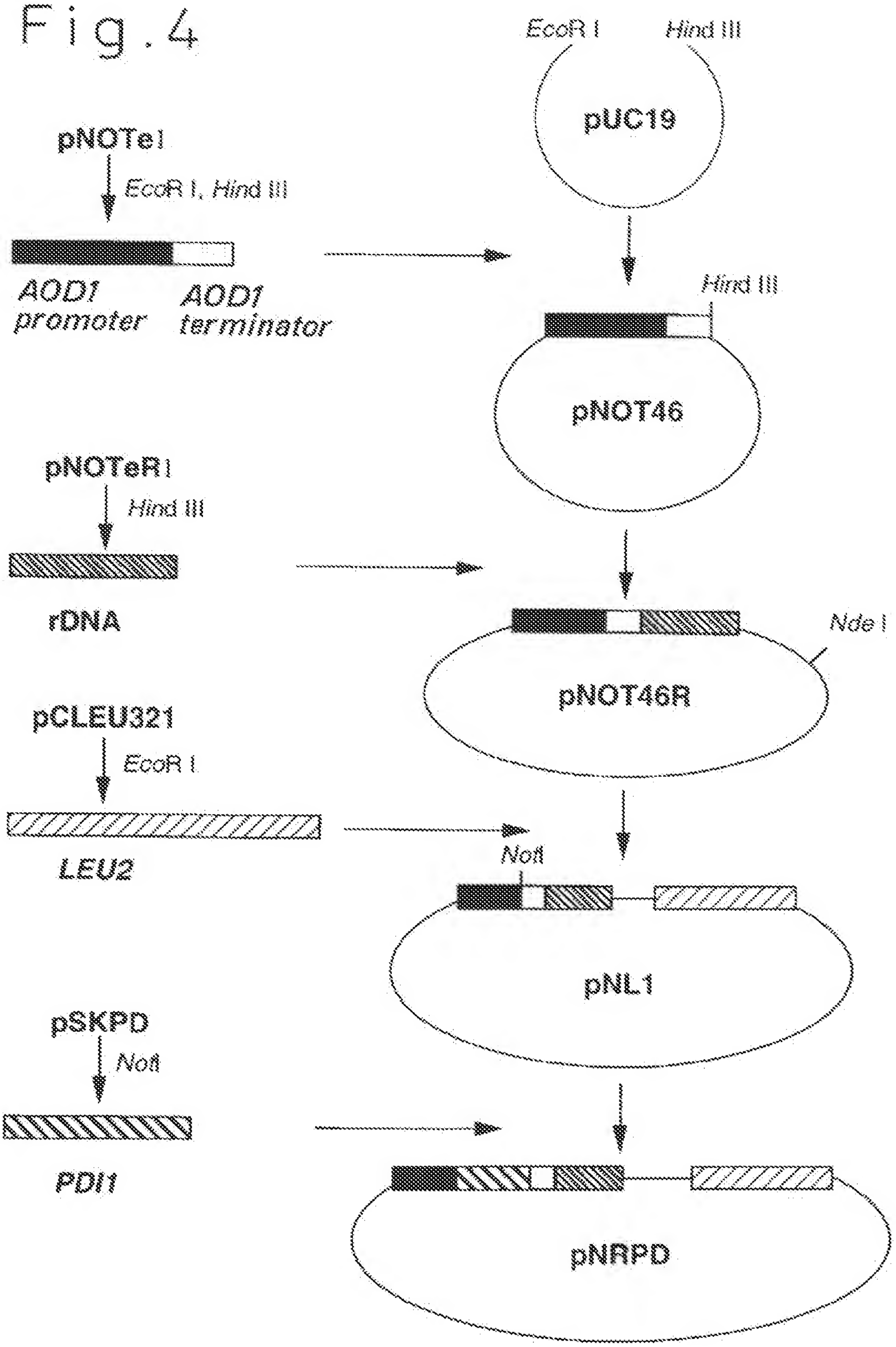

pNOTe1
EcoR I, Hind III
AOD1 promoter
AOD1 terminator
EcoR I
Hind III
pUC19
Hind III
pNOT46
pNOTeR1
Hind III
rDNA
pNOT46R
Nde I
pCLEU321
EcoR I
LEU2
Not I
pNL1
pSKPD
Not I
PDI1
pNRPD

Fig. 5

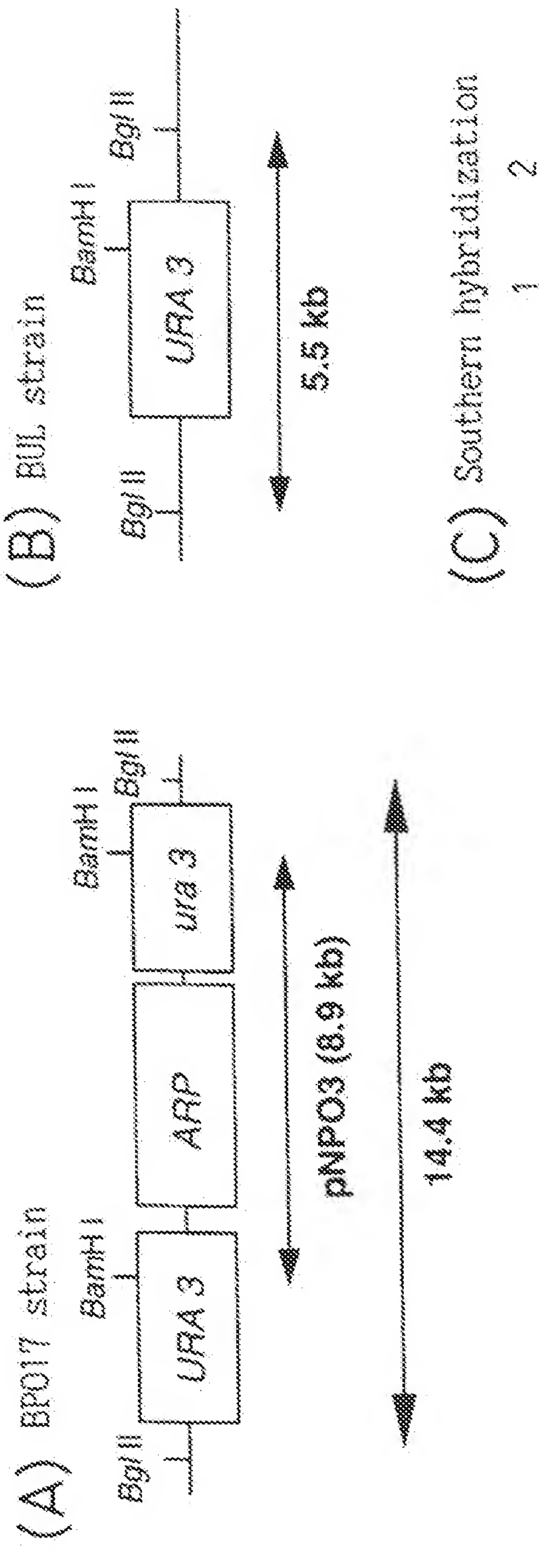

(A) BP017 strain
BamH I
Bgl III
URA 3
Bgl III
BamH I
5.5 kb
(B) BIL strain
Bgl III
BamH I
ura 3
ARP
pNPO3 (8.9 kb)
14.4 kb
(C) Southern hybridization
1
2
14.4kb
5.5kb
1: BP017 strain
2: BIL strain

Fig. 6

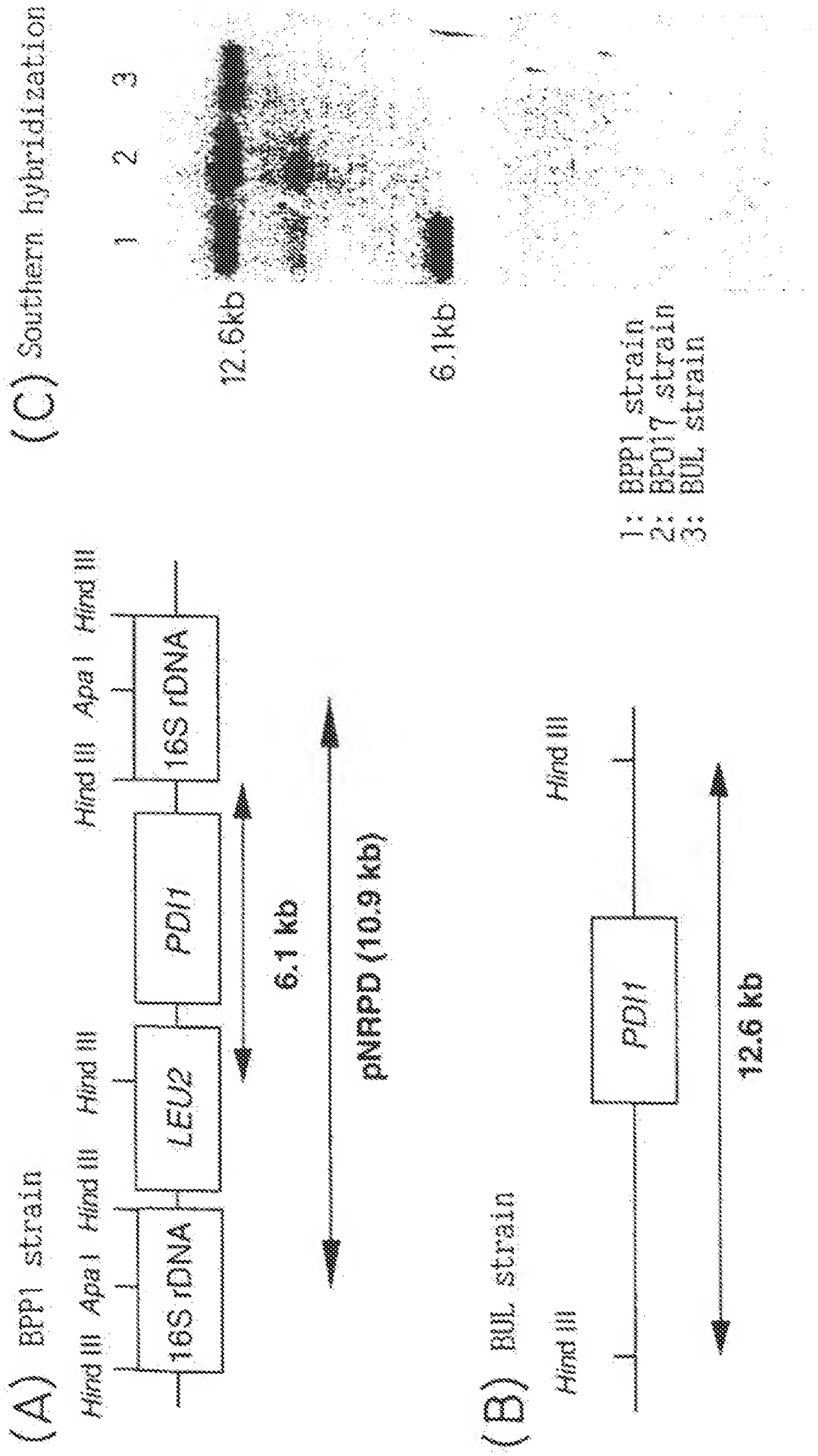
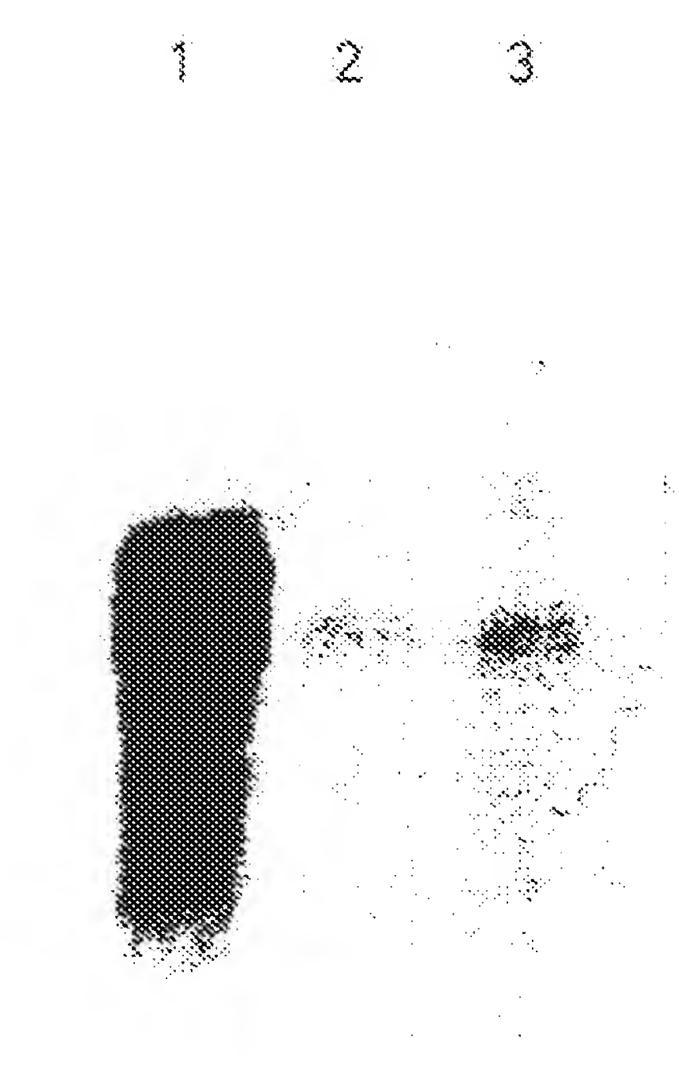
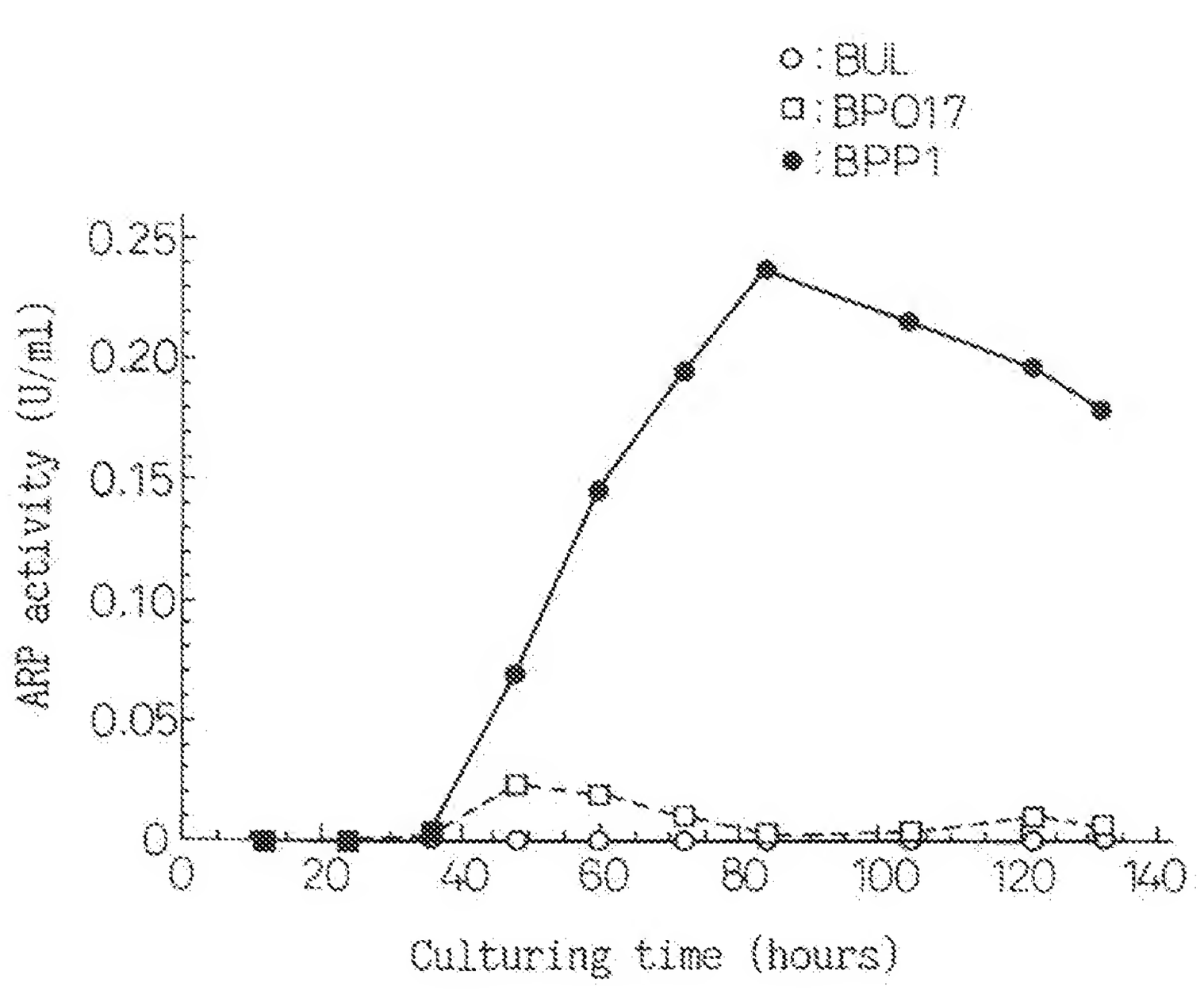

(A) BPP1 strain
Hind III Apa I Hind III
16S rDNA
PDI1
LEU2
Hind III
6.1 kb
pNRPD (10.9 kb)
(B) BUL strain
Hind III
PDI1
Hind III
12.6 kb
(C) Southern hybridization
1 2 3
12.6kb
6.1kb
1: BPP1 strain
2: BPD7 strain
3: BUL strain

Fig. 7



1: BPP1 strain
2: BPO17 strain
3: BUL strain

Fig.8


○ : BUL
□ : BPO17
● : BPP1
ARP activity (U/ml)
0.25
0.20
0.15
0.10
0.05
0
0
20
40
60
80
100
120
140
Culturing time (hours)